The influence of empirical contact networks on modelling diseases in cattle

Duncan, A J ; Gunn, G J ; Lewis, F I ; Umstatter, C ; Humphry, R W

Abstract: We present two stochastic models of the passage of an SEIR (susceptible–latent–infected–resistant) disease through herds of cattle. One model is based on a contact network constructed via continuously recorded interaction data from two herds of cattle, the other, a matching network constructed using the principles of mass-action mixing. The recorded contact data were produced by attaching proximity data loggers to two separate herds of cattle during two separate recording periods. The network constructed using the principles of mass-action mixing uses the same number of contacts as the recorded network but distributes them randomly amongst the animals. The recorded networks had a greater number of repeated contacts, lower closeness and clustering scores and greater average path length than the mass-action networks. A lower proportion of simulations of the recorded network produce any disease spread when compared to those simulations of the mass-action network and, of those that did, fewer infected animals were predicted. For all parameter values tested, within the sensitivity analysis, similar differences were found between the recorded and mass-action network models.

DOI: https://doi.org/10.1016/j.epidem.2012.04.003

The influence of empirical contact networks on modelling diseases in cattle

A. J. Duncan[a,*], G. J. Gunn[b], F. I. Lewis[c], C. Umstatter[d], R. W. Humphry[b]

[a]*Inverness College UHI, Longman Campus, 3 Longman Road, Longman South, Inverness, IV1 1SA*

[b]*Epidemiology Research Unit, SAC (Scottish Agricultural College), Drummondhill, Stratherrick Road, Inverness, IV2 4JZ*

[c]*Vetsuisse Faculty, University of Zurich, Winterthurerstrasse 270, Zurich CH-8057, Switzerland*

[d]*Hill and Mountain Research Centre, SAC (Scottish Agricultural College), Kirkton, Crianlarich, FK20 8RU*

[*]Corresponding author

Email address: andrew.duncan@inverness.uhi.ac.uk (A. J. Duncan)

Abstract

We present two stochastic models of the passage of an SEIR (susceptible–latent–infected–resistant) disease through herds of cattle. One model is based on a contact network constructed via continuously recorded interaction data from two herds of cattle, the other, a matching network constructed using the principles of mass–action mixing. The recorded contact data were produced by attaching proximity data loggers to two separate herds of cattle during two separate recording periods. The network constructed using the principles of mass–action mixing uses the same number of contacts as the recorded network but distributes them randomly amongst the animals. The recorded networks had a greater number of repeated contacts, lower closeness and clustering scores and greater average path length than the mass–action networks. A lower proportion of simulations of the recorded network produce any disease spread when compared to those simulations of the mass–action network and, of those that did, fewer infected animals were predicted. For all parameter values tested, within the sensitivity analysis, similar differences were found between the recorded and mass–action network models.

Keywords: network, mass–action, disease, recorded contacts, SEIR simulation, upland beef

1. Introduction

The assumption of mass–action mixing is prevalent in modelling the transmission of infectious diseases [1, 2]). Despite its widespread use, studies which provide empirical support for the appropriateness, or otherwise, of mass–action mixing are rare [3]. Using data from a study of cattle interactions within their grazing environment, we present results that demonstrate firstly that there are structural differences in the recorded network when compared with a mass–action type network. Secondly, there were differences in the spread of disease modelled through a network of recorded contacts versus the same disease modelled through a network constructed using mass–action mixing. In each of our examples, we found that fewer simulations predicted disease in the recorded model than the mass–action model and of those simulations that predicted disease, fewer animals were infected in the recorded network model than in the mass–action model.

The small number of studies that have empirically examined homogeneous mixing (including both true and pseudo mass–action mixing) within animal populations report varied results. For example, it has been found unsuitable for modelling virus spread in the larvae of the gypsy moth [4], the African army–worm [5] and does not adequately describe the transmission of bovine tuberculosis in bushtail possums [6]. However, it was found suitable for the modelling of brucellosis within herds of bison [7] and pseudorabies in domestic pigs [8].

In addition to the empirical testing of mass–action using observed interactions, this assumption has been explored theoretically through comparison of models with more complex mixing structures involving explicit contact networks. It has been shown that simple homogeneous mixing can successfully model disease transmission in populations where each individual is continuously moving on a straight line path or when the duration of the contacts are much shorter than the period of infection of the disease [9, 10]. However, mass–action mixing needs modification to model disease spread through networks where the contacts are clustered [11] and is unsuitable for modelling the transmission of some sexually transmitted infections [12]. Additionally, the results from a mass–action model were found to be very different from four theoretical network models: clustered; a superspreader; paired contacts and a social network [13].

The appropriateness of mass–action mixing when modelling animal to animal disease transmission within livestock herds remains unassessed. A previous modelling study of the transmission of *E. coli. O157* in a dairy herd [14] compared mixing through an explicit contact network with homogeneous mixing, where the contacts used to construct this network were based on a two hour observation period of the animals involved. Observed contact data has also been used to construct analytic models for disease spread in sheep flocks [15]. In this case the contacts of individual animals were observed over randomly allocated ten minute intervals during daylight hours for fifteen days per flock. These observation periods are both short and non–continuous, so more extensive study of contact interactions may provide more robust results. Our

contact data relates to herds of cattle grazed in upland pastures recorded continuously for periods of 28 to 30 days.

Proximity data loggers [16] were used to record all interactions between animals within each herd. These loggers are worn as collars and when they come within a set distance of one another, log the time and duration of contact along with the other collar's identification number (ID). Proximity loggers have been successfully used previously to study the behaviour of interactions between cows and their calves [17].

In this study, results from two stochastic models are compared. One model uses the network of contacts recorded by the proximity loggers and the other uses a network based on the principle of mass–action mixing. The passage of a hypothetical disease is modelled through these networks to establish whether there is any difference between the results from recorded contacts and contacts created by assuming mass–action. We demonstrate that there are quantifiable differences in the properties of the network of recorded contacts and a mass–action type network and that there are also differences in the spread of the disease between the two models.

2. Materials and Methods

The contact data were recorded by two herds of cattle fitted with proximity data loggers on two different hillside grazing sites.

2.1. Proximity Data Loggers

The contacts in the data sets available were recorded using proximity data loggers [18]. These loggers are worn as collars and are designed specifically to log contacts between individuals. Each logger broadcasts a unique ID on Ultra–High–Frequency bandwidth (UHF). When two loggers come within the proximity distance of each other they begin to log a contact. Both loggers store the opposite ID number, the start time and date of the contact along with the duration.

Before the deployment of the loggers, two important parameters need to be set, the separation time and the UHF Range Coefficient. The separation time is a user defined period. If the contact is broken for longer than that period, the contact is regarded as finished and logged as a single contact. Owing to the nature of UHF, the contact is easily broken if the animals, for example, turn their heads. The UHF Range coefficient determines the proximity threshold.

The separation time was set at 240 seconds to avoid the memories of the loggers reaching capacity before the recording period ended. The loggers were set with a UHF Range coefficient of 50. It has been found that in field trials the UHF setting of 50 recorded an average contact initiation distance of 1.36 ± 0.18m and an average contact termination distance of 2.61 ± 0.23m [19].

2.2. Data Sets

In total, four distinct data sets were available for analysis, recorded by two herds on two upland grazing sites. Both herds comprised 15 cows with their calves at foot. The first herd consisted of Aberdeen Angus cross Limousin cows and the second herd contained purely Luing cows. During the first recording period one bull was present in each herd. The four data sets were recorded in two separate, four week periods. Herd 1 grazed on a 286 ha pasture whilst herd 2 grazed on a 172 ha pasture. All grazing areas were shared with sheep which is typical of Scottish upland grazing sites. No supplementary feeds were used during the recording periods.

To ensure a continuous recording period in each of the datasets any animal that did not register a contact for an entire 24 hour period was removed from the dataset. These missing periods were due to either collar failures, battery failures, mis–recording of data or removal of animals. In the first recording period both bulls were removed from the datasets for this reason. Those animals whose recorded interactions were removed were still present on the hillsides but not represented in the networks.

The first two data sets were recorded during July 2009, both producing 30 complete days of usable data with both of the herds returning complete data for 29 animals. The final two data sets recorded 28 complete days of data across August and September 2009 with herd 1 returning data for 21 animals while herd 2 returned data for 17 animals.

Within this paper we have used the labels dataset 1 and dataset 2 for the datasets produced in July by herds 1 and 2 respectively. The datasets produced in August/September by herds 1 and 2 have been labelled dataset 3 and dataset 4 respectively.

2.3. Network Construction and Comparison

For each of the four datasets a contact network was established, with the nodes representing the animals, and the edges as the recorded interactions. These networks were split into consecutive 12 hour time steps to give a manageable number of edges for each step in the later disease simulation. An identical number of mass–action networks were constructed by taking the total number of interactions recorded in the particular 12 hour period for a particular dataset, creating the same number of random contacts and randomly allocating each of these contacts to pairs of animals in the respective herd. For each dataset and 12 hour period this gave us two networks, a recorded contact network and a mass–action network, with the same number of nodes and edges but with a different distribution of edges. A representative plot of these two types of network can be seen in figure 1.

Although these two types of network contain the same number of nodes (animals) and edges (contacts), it is clear the edges are distributed differently. Upon further investigation the largest difference found in the distribution of the edges, was that in the network of recorded contacts, many of the pairwise

combinations of connected nodes were repeated. To quantify such structural differences between the two types of networks, four network metrics were used. The first was our own metric, the number of repeated edges, chosen to quantify the observed difference exemplified in figure 1. The second was closeness, the inverse of the average length of the shortest paths to/from all the other vertices in the network [20], and the third metric chosen was the clustering coefficient, a measure of the degree to which nodes in a network tend to cluster together [21]. The fourth and final metric that we used, average path length [22], is the average number of steps along the shortest path for all possible pairs of nodes. These metrics were calculated for each type of network, for each 12 hour period and for all four datasets. To take into account the random allocation of edges between nodes in the mass–action networks, 1000 mass–action networks were constructed for each 12 hour period in each dataset. To make it easier to establish any differences between the two networks the mean differences in the metrics were found by averaging the differences over all 12 hour periods in the dataset. For each dataset and each metric the mean difference is given explicitly as

$$\text{Mean Difference} = \frac{\sum_{t=1}^{n} ({}^{mass}m_t - {}^{rec}m_t)}{n}, \tag{1}$$

where n represents the number of 12 hour time periods in the dataset, the superscripts rec and $mass$ denote the results from the recorded contact networks and mass–action networks respectively and m denotes the median value for the metric. This value of m is calculated in two ways. The metric of

closeness returns values for each node in the network and so m is found via

$$^{mass}m = \underset{\text{over 1000 networks}}{\text{median}} \left(\underset{\text{over all nodes}}{\text{median}} \left(^{mass}\text{closeness} \right) \right) \tag{2}$$

$$^{rec}m = \underset{\text{over all nodes}}{\text{median}} \left(^{rec}\text{closeness} \right). \tag{3}$$

For the other metrics a single value is returned for the network as a whole giving us

$$^{mass}m = \underset{\text{over 1000 networks}}{\text{median}} \left(^{mass}\text{metric} \right) \tag{4}$$

$$^{rec}m = ^{rec}\text{ metric}. \tag{5}$$

2.4. Models and Disease

We use an SEIR [1] type model in our analysis, this is a very flexible and generic modelling framework appropriate for a number of pathogens relevant to cattle health. S, E, I and R denote the number of animals in the susceptible, latent, infectious and resistant classes respectively. An SEIR model can be described by the system of ordinary differential equations (ODEs),

$$\begin{aligned} \frac{dS}{dt} &= -\alpha\beta\frac{SI}{N}, \\ \frac{dE}{dt} &= \alpha\beta\frac{SI}{N} - \sigma E, \\ \frac{dI}{dt} &= \sigma E - \gamma I \\ \text{and} \qquad \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{6}$$

with $S + E + I + R = N$, where N is the total (constant) population size. Each susceptible animal moves from the susceptible state to the latent state with rate $\alpha\beta$ following a contact with an infectious animal, where α is the probability of infection from a single contact with an infectious animal and β is the average number of daily contacts per animal. The parameter σ is the rate at which

those in the latent class move to the infectious class and γ the rate at which animals move from the infectious class to the resistant class. For the SEIR model (6), the basic reproductive number, R_0, is given by

$$R_0 = \frac{\alpha\beta}{\gamma}, \tag{7}$$

[23]. R_0 determines the behaviour of the continuous, deterministic system in (6).

We employed two stochastic versions of (6). The first is based on the recorded contact networks and the second uses the mass–action networks. These models are implemented as stochastic due to the small numbers of animals in each of the data sets, and hence the increased influence of individual stochastic events on the overall disease transmission process [2]. Infection was always introduced by randomly infecting a single animal at the start of each model simulation, thus this animal began the simulation in the latent state. The period of time each animal spends in the latent and infectious states were sampled from exponential distributions with means $1/\sigma$ and $1/\gamma$. For simplicity, and because the largest dataset only contained 30 days of continuously recorded interactions, each infected animal had its length of resistance set to greater than 30 days. Both models were simulated many times and it was found that the probability densities of the number of animals in each disease state at each time point, appeared to stabilise by 5000 simulations. All results presented were produced from 5000 simulations, where each simulation was run for the number of days contained in the respective dataset with an initially infected animal randomly chosen for each simulation and a randomly allocated edge distribution in the

mass–action network.

The value of β, the mean contact rate, used in the simulations was dependent on the data set used, as each of the four data sets had a different average contact rate. Thus we had four values for β corresponding to our four datasets.

The disease spread through each model was a hypothetical disease with parameter values that allowed the peak of infection of an epidemic to occur within the 28 days of data available from the shortest dataset. Latent and infectious periods of six days were chosen. Using average values of $\beta = 7.987$ from our data and $R_0 = 5$ (considered reasonable), a rounded value of $\alpha = 0.1$ was calculated from (7). As each dataset has a different value of β, the contact data, they will also have a different value of R_0 but the characteristics specific to the disease ($\alpha = 0.1$, $1/\sigma = 6$ days and $1/\gamma = 6$ days) remain fixed across all datasets. It was assumed that when an animal became infected its behaviour did not change such that its contacts continued as normal. This is not necessarily the case [24, 25] but until there exists actual contact data for a herd with spreading disease, it is parsimonious to use the actual data that we do have.

3. Results

3.1. Network Comparison

To establish the difference between the two types of network used in the models, four metrics were calculated: average path length, closeness, clustering coefficient and the number of repeated edges. Table 1 shows the values of these

metrics for each dataset along with the mean differences (with 95% confidence intervals on the mean differences calculated over all 12 hour periods) calculated using (1).

The biggest difference is found in the number of repeated edges. The recorded networks have more repeated edges and slightly longer average path lengths. The mass–action networks have slightly higher clustering coefficients and higher values of closeness. It is interesting to note that for the fourth dataset, a difference of zero is contained with the confidence intervals shown for both the average path length and clustering coefficient, yet there are still, on average, 21 more repeated edges in the network of recorded contacts. As can be seen in table 1, this figure represents a more than 100% increase in the number of repeated edges from the mass–action networks to the networks of recorded contacts.

It seems possible that the differences in the other metrics are largely influenced by the differences in the number of repeated contacts. Examination of where repeated contacts occurred showed 32%, 22%, 18% and 15% of all the repeated contacts in datasets 1, 2, 3 and 4 respectively were between a mother and her calf. If repeated contacts were equally likely between any two animals irrespective of their relationship then the expected proportion of repeated contacts that are mother–calf would be $1/(a-1)$, where a signifies the number of animals in the herd. This would give a figure of 3-4% for our herds. It therefore appears that a significant contributor to the structural network differences between the recorded networks and those constructed using

mass–action mixing is the strong mother–calf contact. This is in keeping with other studies using proximity collars to monitor animal movement and within–herd interaction [17, 26].

3.2. Disease Spread

Figure 2 shows the proportion of the 5000 simulations that had no infected animals through the recording periods. The results for the recorded model (dashed lines) are shown along with the mass–action model (solid lines) for all four datasets for the standard set of disease parameters. It is clear that a large proportion of the simulations from both models produce no infection in any of the datasets. This results in a bimodal distribution in the number of animals infected and so the simulations that produced disease were separated from those that did not. Figure 3 shows the 25^{th}, 50^{th} and 75^{th} percentiles of the number of infected animals for both the recorded model (black, dashed lines) and mass–action model (red, solid lines). The results are shown for every second time step for clarity.

A sensitivity analysis was conducted on four disease parameters: α; $1/\sigma$; $1/\gamma$; the length of time period used to construct the network, along with the period of time considered to be a contact and the amount of missing data. Each of the disease parameters were individually halved and doubled. Due to settings on the proximity collars the four minute contact length could not be halved so contact lengths of 8 and 16 minutes were utilised as alternative definitions of a contact.

To assess the sensitivity to missing data, three other scenarios were considered. Firstly, an additional animal was removed at random from each dataset. This was done six times to see if there was any difference caused by the particular animal removed. Secondly, in one dataset there was an animal whose collar stopped recording two days before the end of the recording period and its data was re–added to reflect a relaxation in the construction of the datasets. In our third scenario assessing sensitivity to missing data, any animal with no contacts for any 12 hour period had that period filled with contacts by randomly sampling one of the 12 hour periods where the animal did record contacts. For this third scenario, to reflect as complete datasets as possible, any animal not in the original datasets but with more than five days of recording contacts was added back to the datasets with their missing days replaced as above. This process was carried out six times to establish if there were any differences from the sampling process.

Our primary interest in the sensitivity analysis is to assess the sensitivity of the difference between the recorded and mass–action networks. Accordingly we examine both the sensitivity of the disease spread and the sensitivity of the difference between the recorded and mass–action models in turn.

3.2.1. Sensitivity of Disease Spread

The largest sensitivity in disease spread occurred when the contact length was increased. This was the case for both the proportion of 5000 simulations that produce infection and the number of infected animals predicted by those

simulations. Other high sensitivity in disease spread was found when the ratio of $1/\sigma$:$1/\gamma$ was 2:1. When the contact lengths of 8 and 16 minutes were used, far fewer contacts were recorded. Plots for the 8 minute contact length and for the case when the ratio of $1/\sigma$:$1/\gamma$ was 2:1 can be seen in the online supplementary information.

3.2.2. Sensitivity of Difference in Spread of Disease between recorded and mass–action models

The differences between the recorded and mass–action models were not sensitive to any of the parameters. In those simulations that produced infection there was a consistent mean absolute difference of between 0.3 and 0.9 in the number of infected animals. The only time this difference was greater than one animal, occurred when $1/\sigma$:$1/\gamma$ was 2:1 in dataset 1, this represented approximately 3% of the herd. Full results are presented in tables 1 and 2 in the online supplementary information.

3.2.3. Overall Comparison of Differences between recorded and mass–action models

The main aim of this work was to establish if there was a difference in the spread of the disease through the recorded network when compared to the mass–action network. The results of figures 2 and 3 along with those in the supplementary information show that fewer simulations produce infection in the recorded network and those that do produce infection have less infected animals than the mass–action network. This is shown in greater detail in the online supplementary information. The mean differences in the proportion of runs,

produced by each model, which predict infection are always positive showing that the recorded network model always produced fewer simulations showing infection. Furthermore, the mean differences in the number of infected animals, predicted by each model, are always negative. This demonstrates that the recorded network model always predicted fewer infected animals, consistently over all the sensitivity scenarios.

4. Discussion

The results show that there are differences in the properties of the recorded and mass–action networks with the networks constructed from recorded contact data having more repeated edges, lower closeness and clustering and higher average path lengths. Datasets 3 and 4 (from the second recording period) had noticeably fewer repeated edges than datasets 1 and 2. This, in part, almost surely reflects the smaller number of animals returning valid data. Other concomitant biological changes between the two recording periods are the decreasing calf–cow bond with age, the presence of a bull in the first recording period but not the second and any other miscellaneous changes (e.g. weather, grazing quality).

The passage of the hypothetical disease has been shown to be different through the recorded and mass–action networks. The distributions of the number of infected animals produced by the 5000 simulations of each network were bimodal with a large proportion of these producing no infection. A greater proportion of the simulations from the recorded network produced no infection

and of those simulations that did produce infection there were on average fewer infected animals compared to the mass–action network. This was the case for all four datasets for almost all of the variations in the sensitivity analysis.

Within the sensitivity analysis the largest differences in disease spread, occurred when the contact length was increased to 8 or 16 minutes. This brings up the question of what is considered to be a contact. The proximity data used within this work only gives us an idea of location of the animals with respect to one another, not what activity they are engaged in. Different types of contact have been used in the modelling of human diseases, where questionnaires have been utilised to establish the length and strength of the contacts [27, 28, 29, 30]. In modelling the diseases of animals this is harder to achieve but observation has been used in the past [15]. These observation periods were short when compared to the continuous data that can be gathered from the collars suggesting possibly an amalgamation of the two recording methods would provide an improved model. However, regardless of what is considered to be a contact (4, 8 or 16 minutes) the recorded model produced fewer simulations which result in infection. In those simulations with infection, fewer animals are infected when compared to the mass–action model. Whilst the disease spread is sensitive to contact length, it is comforting that the resultant difference between spread in a recorded network and in a mass–action was not.

Whether these sensitivities would appear if the characteristics of the herd were different (e.g. larger herd size, different time of year, housed, or older calves) is

something we cannot comment on with the data available to us. However, we hope that these results encourage additional investigation of such factors.

The model disease's parameters $\alpha = 0.1$, $1/\sigma = 6$ days and $1/\gamma = 6$ days were chosen to ensure that the peak of infection of the epidemic could take place in the 28 days of recorded data available. These parameters are not merely hypothetical and convenient for the data we had - they approximate those published for two diseases of cattle: foot and mouth ($R_0 = 2.1 - 73$, $1/\sigma = 2.5$ days, $1/\gamma = 5$ days [31, 32]) and transient bovine viral diarrhoea (BVD) ($R_0 = 2.3$, $1/\sigma = 3$ days, $1/\gamma = 10$ days [33, 34]). The model disease is not a predictive representation of any of these three diseases because they are not purely spread by direct animal to animal to contact, the only transmission method that can be modelled with the proximity data available.

Whilst noting that the inherent structure of our real contact data does influence the course of the simulated epidemic, it is worth also noting the high levels of uncertainty in the literature estimates of the basic disease parameters. One possible consequence of this is that any increase in precision obtained in models by accounting for true network structure will only be matched by an increase in accuracy when the other disease parameters are better known.

5. Conclusion

We have shown that there is a difference in the network properties of a network constructed from recorded contacts when compared to a mass–action type

network. We find corresponding differences in predicted spread of an SEIR type disease when animal contacts are taken from recorded data versus constructed by the mass–action mixing. Through a sensitivity analysis we have found that both the recorded network model and the mass–action model are most sensitive to the definition of a contact. However, similar differences were found between the two models, for all parameter values. Fewer simulations of the recorded network produce any disease spread when compared to those simulations of the mass–action network and of those that did, fewer infected animals were predicted. The differences found lead the authors to believe that the network structure needs to be considered when modelling diseases of cattle spread via direct animal to animal contact.

6. Acknowledgements

The Scottish Agricultural College receives financial support from the Scottish Government. The authors would like to acknowledge the link between Inverness College UHI and the Scottish Agricultural College thus enabling this work to be carried out. Finally, the authors would like to acknowledge the comments on the reviewers which led to very salient additional work to be carried out.

References

[1] R. Anderson, R. May, Infectious Diseases of Humans, Oxford University Press, Oxford, UK, 1991.

[2] F. Brauer, P.W.J. van den Driessche (Eds)., Mathematical Epidemiology, Lecture Notes in Mathematics, Springer, 2000.

[3] H. McCallum, N. Barlow, J. Hone, How should pathogen transmission be modelled? Trends in Ecol. & Evol. 16 (2001) 295–300. doi:10.1016/S0169–5347(01)02144–9.

[4] V. D'Amico, J.S. Elkinton, G. Dwyer, J.P. Burand, J.P. Buonaccorsi, Virus transmission in gypsy moths is not a simple mass action process, Ecol. 77 (1996) 201–206.

[5] A.F. Reeson, K. Wilson, J.S. Cory, P. Hankard, J.M. Weeks, D. Goulson, R.S. Hails, Effects of phenotype plasticity on pathogen transmission in the field in a lepidoptera–npv system, Oecologia 124 (2000) 373–380.

[6] N.D. Barlow, Non-linear transmission and simple models for bovine tuberculosis. J. Anim. Ecol. 69 (2000) 703–713.

[7] A. Dobson, M. Meagher, The population dynamics of brucellosis in the yellowstone national park, Ecol. 77 (1996) 1026–1036.

[8] A. Bouma, M.C.M. Dejong, T.G. Kimman, Transmission of psuedorabies virus within pig–populations is independent of the size of the population, Prev. Vet. Med. 23 (1995) 163-172. doi: 10.1016/01675877(94)00442–L.

[9] C.J. Rhodes, R.M. Anderson, Contact rate calculation for the basic epidemic model, Math. Biosci. 216 (2008) 56–62. doi:10.1016/j.mbs.2008.08.007.

[10] E. Volz, L.A. Meyers, Susceptible–infected–recovered epidemics in dynamic contact networks, Proc. Roy. Soc. B 274 (2007) 2925–2933.

[11] M. Keeling, The implications of network structure for epidemic dynamics, Theor. Pop. Biol. 67 (2005) 1–8. doi: 10.1016/j.tpb.2004.08.002.

[12] R.M. May, R.M. Anderson, Transmission dynamics of hiv–infection, Nature 326 (1987) 137–142.

[13] C.R. Dietrich, H.H. Thulke, Influence of contact structure between host animals on the dynamics of infectious diseases, in: T.D.H. Parkin, L.A. Kelly (Eds.), Proceedings of Society for Veterinary Epidemiology and Preventative Medicine Leipzig: Society for Veterinary Epidemiology and Preventative Medicine, 2012, pp. 157–169.

[14] J. Turner, R.G. Bowers, D. Clancy, M.C. Behnke, R.M. Christley, A network model of *E.coli O157* transmission within a typical UK dairy herd: The effect of heterogeneity and clustering on the prevalence of infection, J. Theor. Biol. 254 (2008) 45–54. doi:10.1016/j.jtbi.2008.05.007.

[15] D. Schley, E. Norton, S. Whittle, S. Benaben, D.R.N. Mbotha, M. Taylor, I.Z. Kiss, Contact networks amongst domestic sheep: the potential for disease transmission in flocks, in: C. Fourichon, D.U. Pfeiffer (Eds.), Proceedings of Society for Veterinary Epidemiology and Preventative Medicine Leipzig: Society for Veterinary Epidemiology and Preventative Medicine, 2011, pp. 182–192.

[16] Sirtrack Proximity Collar Technical Details. http://www.sirtrack.com/documents/suites/Proximity.pdf

[17] D.L. Swain, G.J. Bishop–Hurley, Using contact logging devices to examine

animal affiliations: quantifying cow–calf interactions, Appl. Anim. Behav. Sci. 102 (2007) 1-11. doi:10.1016/j.applanim.2006.03.008.

[18] S. Prange, T. Jordan, C. Hunter, S.D. Gehrt, New radiocollars for the detection of proximity among individuals, Wildl. Soc. Bull. 34 (2006) 1333-1344.

[19] E.L. Goodman, Quantifying interactions in a high–density badger (*Meles Meles*) population (PhD thesis), York, UK, University of York, 2007.

[20] G. Csardi, Package 'igraph', last updated 18/10/2011, http://cran.r-project.org/web/packages/igraph/igraph.pdf, last accessed 24/10/2011.

[21] M.E.J. Newman, The Structure and Function of Complex Networks, SIAM Rev. 45 (2003) 167–256.

[22] S.H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.

[23] M. Li, J. Muldowney, P. van den Driessche, Global stability of seirs models in epidemiology, Can. Appl. Math. Q. 7 (1999) 409–425.

[24] J.W. Wilesmith, Manual on bovine spongiform encephalopathy, FAO Animal Health Manual 2, FAO, Rome, Italy, 1998, pp. 34.

[25] J. Rush, A.M. de Passille, M.A.G. von Keyserlingk, D.M. Weary, The Welfare of Cattle Volume 5, Springer, Dordrecht, The Netherlands, 2008, pp. 37.

[26] R.N. Handcock, D.L. Swain, G.J. Bishop–Hurley, K.P. Patison, T. Wark, P. Valencia, P. Corke, C.J. O'Neill, Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing, Sensors 9 (2009) 3586–3603. doi:10.3390/s90503586.

[27] K.T.D. Eames, J.M. Read, W.J. Edmunds, Epidemic prediction and control in weighted networks, Epidemics 1 (2009) 70–76.

[28] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G.S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinka, W.J. Edmunds, Social contacts and mixing patterns relevant to the spread of infectious diseases, PLOS Med. 5 (2008) 381–391.

[29] J.M. Read, K.T.D. Eames, W.J. Edmunds, Dynamic social networks and the implications for the spread of infectious disease, J. Roy. Soc. Interface 5 (2008) 1001–1007.

[30] J. Wallinga, P. Teunis, M. Kretzschmar, Using data of social contacts to estimate age–specific transmission parameters for respiratory–spread infectious agents, Am. J. Epidemiol. 164 (2006) 936–944.

[31] M.E.J. Woolhouse, D.T. Haydon, A. Pearson, R.P. Kitching, Failure of vaccination to prevent outbreaks of foot–and–mouth disease, Epidemiol. Infect. 116 (1996) 363–371.

[32] R.P. Kitching, D.K. Mackay, Foot and Mouth Disease, State Vet. J. 4 (1994) 7–10.

[33] B.R. Cherry, M.J. Reeves, G. Smith, Evaluation of bovine viral diarrhea virus control using a mathematical model of infection dynamics, Prev. Vet. Med. 33 (1998) 91–108.

[34] J.S. Nickell, B.J. White, R.L. Larson, D.G. Renter, J. Roque, R. Hesse, R. Oberst, L. Peddireddi, G. Anderson, Onset and duration of transient infections among antibody–diverse beef calves exposed to a bovine–viral diarrhea virus persistently infected calf, Int. J. Appl. Res. Vet. Med. 9 (2011) 29–39.

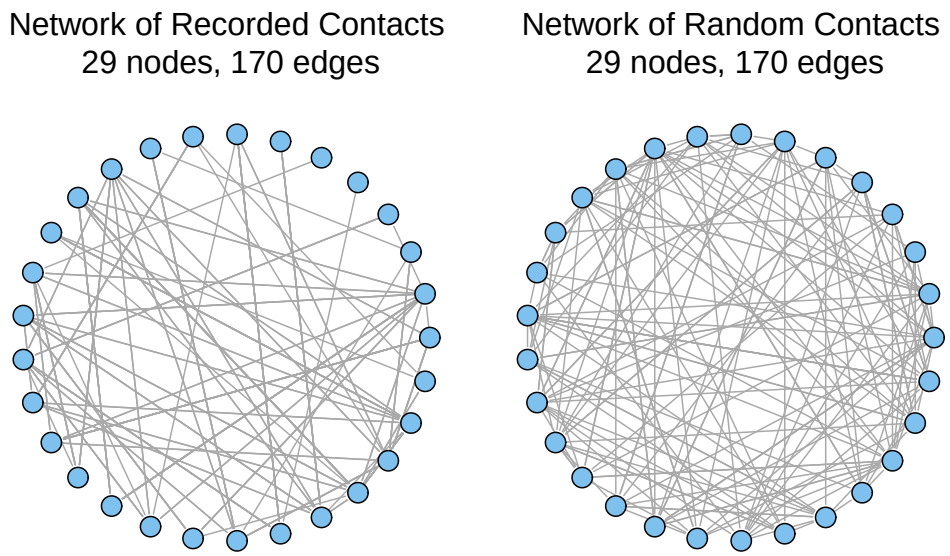

Figure 1: Plot of two 12 hour networks, both constructed with 29 nodes and 170 edges. One constructed with the recorded data, the other using the mass–action assumption of random contact placement. These plots represent the contacts from the 21^{st} 12 hour period of dataset 1.

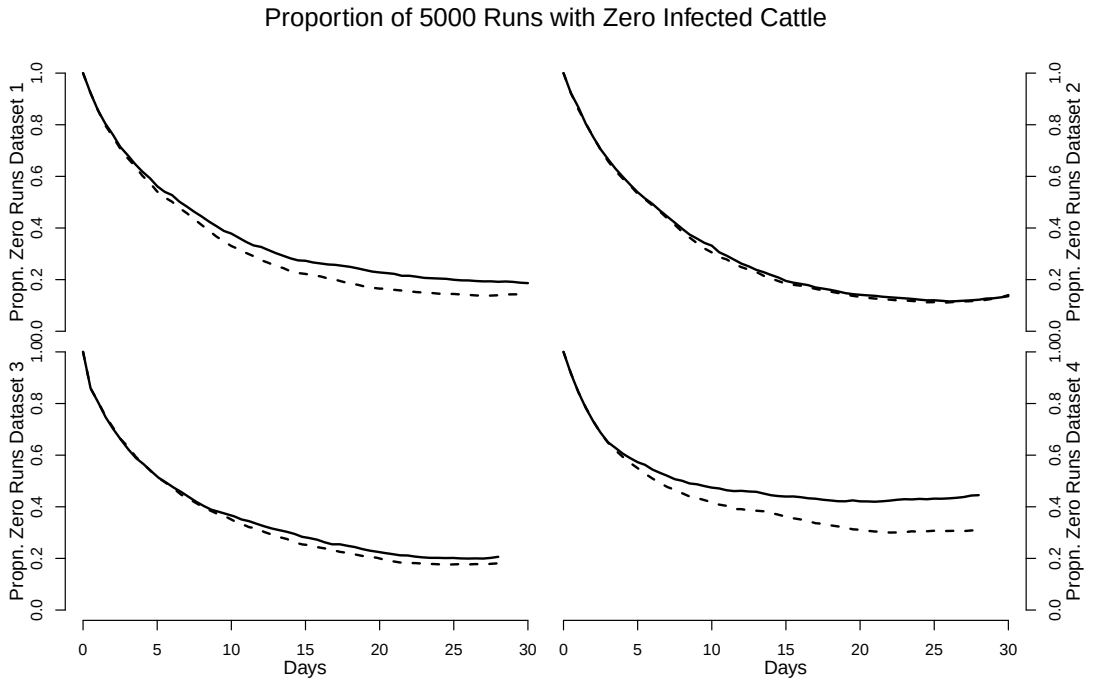

Figure 2: Plot of proportion of 5000 simulations which resulted in no infected animals for each day of the simulation. Results from the recorded model (dashed lines) are shown alongside results from the mass–action model (solid lines) for all four datasets with the standard set of disease parameters $\alpha = 0.1$, $1/\sigma = 6$ days and $1/\gamma = 6$ days.

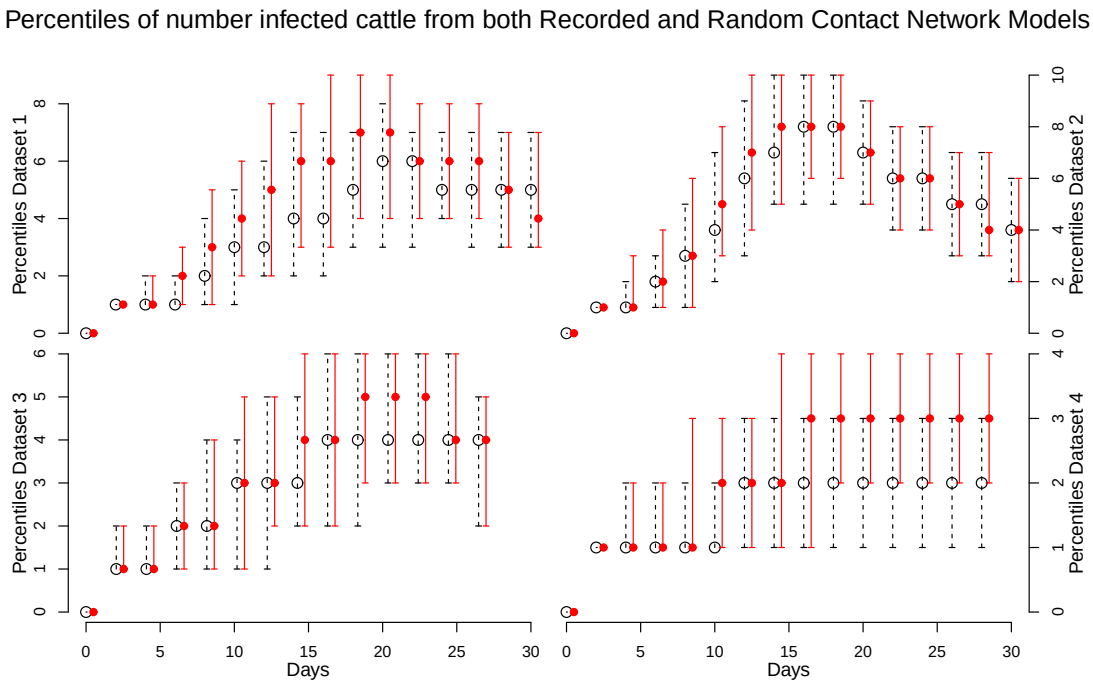

Figure 3: Plot of 25^{th}, 50^{th} and 75^{th} percentiles of the number of infected animals in those simulations which did result in infected being present. Results from the recorded model (black, dashed lines) are shown alongside results from the mass–action model (red, solid lines) for all four datasets with the standard set of disease parameters $\alpha = 0.1$, $1/\sigma = 6$ days and $1/\gamma = 6$ days.

Table 1: Values of medians for both recorded networds and mass–action networks, along with the mean difference (with 95% confidence intervals of the mean difference) of the two values averaged over all 12 hour time periods. The values in the table were calculated using equations (1) to (5). Datasets 1 and 2 were produced in July by herds 1 and 2 respectively. Datasets 3 and 4 were produced in August/September by herds 1 and 2 respectively.

	Metrics			
	Number Repeated Edges	Closeness	Clustering	Average Path Length
Dataset 1				
${}^{mass}m$	62.7583	0.5427	0.3157	1.9056
${}^{rec}m$	141.6	0.2474	0.2909	2.4369
Mean Difference	-78.8417	0.2952	0.0247	-0.5313
95% C.I.	(-85.9341,-71.7492)	(0.2625,0.3280)	(0.0024,0.0471)	(-0.6584,-0.4043)
Dataset 2				
${}^{mass}m$	139.8167	0.6245	0.4407	1.6589
${}^{rec}m$	215.6	0.4423	0.3549	2.0156
Mean Difference	-75.7833	0.1822	0.0858	-0.3567
95% C.I.	(-81.2117,-70.3550)	(0.1585,0.2059)	(0.0660,0.1056)	(-0.4244,-0.2890)
Dataset 3				
${}^{mass}m$	40.3455	0.5259	0.3178	1.8736
${}^{rec}m$	71.3636	0.3450	0.2579	2.3133
Mean Difference	-31.0182	0.1809	0.0598	-0.4397
95% C.I.	(-34.3803,-27.6561)	(0.1539,0.2079)	(0.0389,0.0807)	(-0.5330,-0.3464)
Dataset 4				
${}^{mass}m$	18.6696	0.4443	0.2434	2.0889
${}^{rec}m$	40.0179	0.1453	0.2885	2.1071
Mean Difference	-21.3482	0.2990	-0.0407	-0.0181
95% C.I.	(-23.7111,-18.9853)	(0.2628,0.3353)	(-0.0859,0.0044)	(-0.2045,0.1682)